

# Mercoledì 19 febbraio 2014 Ore 11.00

Ed. U7 - 4° piano

Via Bicocca degli Arcimboldi, 8 20126, Milano

## Brian Francis

Professor of Social Statistics, Lancaster University (UK)

### The need for outlier detection in latent class analysis – an example from criminal careers research

Over 30 years ago, Aitkin and Tunnicliffe-Wilson's *Technometrics* paper of 1980 provided an interesting application of mixture models, using mixtures of Normal distributions as a method of testing for outliers. On small datasets with multiple start points and a sufficient number of mixture components, mixture modelling can be used to detect outliers. This paper takes the *Technometrics* paper as a starting point to investigate outlier detection in mixture models where there may be multiple outliers. Methods in the literature for continuous data have grown and now include fitting mixtures of t-distributions rather than normal distributions (Peel and McLachlan, 2000) or robustifying through modifying the likelihood in various ways (Fujisawa and Eguchi, 2006; McLachlan and Basford, 1988). The focus of this work in contrast is on discrete data and considers latent class models with a large number of binary or multinomial indicators. In large data sets and with a large number of indicators, poor identification of the underlying structure can easily occur.

Moreover, outliers may not be detectable through classic latent class analysis, and classes will be contaminated. We illustrate this problem with a latent class analysis of the criminal typologies of female offenders in England and Wales. Alternative methods of detecting outliers in such problems will be presented and discussed, and the benefit of developing robust approaches to latent class analysis will be emphasised.



# SEMINARIO

Dipartimento di Statistica e Metodi Quantitativi